BILLBOARD AND CINEMA ADVERTISING:

MISSED OPPORTUNITY OR SPOILED ARMS?

Steffi Frison[a,*]

Marnik G. Dekimpe[b,a,i]

Christophe Croux[a,ii]

Peter De Maeyer[c,iii]

[a] KU Leuven – University of Leuven, Naamsestraat 69, 3000 Leuven, Belgium

[b] Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

[c] Mahidol University, 69 Vipawadee Rangsit Road, Bangkok, Thailand 10400

---

[*] Corresponding author. Tel.: +32 16 323 230; fax: +32 16 326 624.
Email addresses: steffi.frison@kuleuven.be (S. Frison), m.g.dekimpe@uvt.nl (M.G. Dekimpe), christophe.croux@kuleuven.be (C. Croux), peter.dem@mahidol.ac.th (P. De Maeyer).

[i] Tel.: +31 13 466 3423; fax: +31 13 466 8354.
[ii] Tel.: +32 16 326 958; fax: +32 16 326 732.
[iii] Tel.: +662 206 2000; fax: +662 206 2090.

Part of the paper was written while the second author was the Tommy Goh Visiting Professor in Entrepreneurship and Business at Singapore Management University. The authors thank GfK Belgium and the Belgian Centre for Information about Media for making the data available, and are indebted to Anirban Mukherjee, Kapil Tuli, Linda Van de Gucht and Harald van Heerde for useful comments on an earlier version of the paper.

ABSTRACT

Advertising remains one of the most popular marketing instruments, and many studies have studied its sales effectiveness. However, prior research has either looked at the total spending of a brand/firm, or has focused on the most popular media, especially TV advertising. Even though huge amounts are also spent on "smaller" media such as billboards and cinema, little is known on their effectiveness.

While many brands never use them (which could be a missed opportunity), others allocate a substantial part of their advertising budget to these media (which could represent spoiled arms in case of insufficient effectiveness). In this study, we conduct a large-scale empirical investigation, using close to seven years of monthly data on over 250 brands of consumer packaged goods, to quantify the sales elasticity of these often-neglected media.

Even though a significant long-run elasticity is found for a number of brands, we obtain a substantially lower proportion of significant effects for billboard and cinema advertising than for the more popular TV medium. Also meta-analytically, and after correcting for the brands' self-selection of media on which to spend their advertising budget, no evidence of a significant short- or long-run sales elasticity is found for these two media, while significant effects are obtained for both TV and magazine advertising. In addition, little evidence of systematic synergy effects with the more traditional media is found. Hence, from a sales-response point of view, investments in billboard and cinema advertising appear to act as spoiled arms for most mature CPG brands.

Keywords: Media usage; Sales response; Advertising elasticity; Billboard advertising; Cinema advertising.

1. INTRODUCTION

Advertising remains one of the most important marketing-mix instruments. Total global adspend in 2012 has been estimated at nearly US$ 500 billion, or 0.7% of the world's global GDP (Barnard 2012). At the company level, advertising represents a substantial investment as well. A recent compilation across more than 5,000 companies from 300+ industries reports an average advertising-to-sales ratio of 3.1% (AdAge 2012). Within the consumer packaged goods (CPG) sector, ratios close to 10% (food products: 9.2%) or higher (e.g., soaps and detergents: 12.1%; cosmetics: 20%) are not uncommon. In an effort to justify these expenditures, a large literature has emerged that quantifies the effectiveness of advertising in terms of its sales elasticity. Recent reviews include, among others, Allenby and Hanssens (2005) and Sethuraman, Tellis and Briesch (2011). The latter compiled a meta-analytic data set of 751 (402) short-run (long-run) advertising elasticities from 56 (38) publications, and report an average elasticity of 0.12 (0.24). Within a CPG setting, van Heerde et al. (2013) studied advertising's sales effectiveness across 150 brands in the U.K., and report average short- and long-run elasticities of 0.002 and 0.013. Srinivasan, Vanhuele and Pauwels (2010), looking at 74 brands across four categories in France, obtained values of 0.020 and 0.036, respectively.

The aforementioned numbers do not recognize, however, that different media (television, radio, magazines, newspapers, ...) can differ widely in their short- and/or long-run effectiveness. Still, that information is critical when deciding on brands' budget allocation across different media (Fisher et al. 2011). To that extent, Sethuraman et al. (2011) distinguished in their meta-analysis between studies reporting, respectively, elasticities for (i) television, (ii) print media, and (iii) an aggregation across multiple media. After accounting for a variety of other factors, they found a significantly higher long-run elasticity for print advertising than for television advertising. For the short-run elasticity, this order was reversed. Other media were not considered separately, given the more limited number of studies that report their elasticities.

Recently, some studies have started to look at a wider variety of media. Naik and Peters (2009), for example, analyzed an advertising campaign for cars involving six media (television, magazines, newspapers, radio, internet banners and sponsored search). They found radio advertising to be most cost effective, followed by newspapers, TV and magazines. Danaher and Dagger (2013), in turn, report the short-run elasticities for ten media in the context of a blitz (one-month) media campaign by an up-market Australian department store. However, given (i) the limited number of such studies, and (ii) their rather unique character (e.g., the blitz-advertising setting in the latter study acted more like a sales-promotion tool, making potential carryover or long-run effects less relevant), no empirical generalizations on those other media are available yet.

This is especially the case for some of the so-called "smaller" media, such as outdoor (billboard) and cinema advertising, which are typically excluded from consideration. Deleersnyder et al. (2009), for example, report across 37 countries the proportion of total advertising spent on four "key media" (p. 628), television, radio, magazines and newspapers, for which they analyze the cyclical sensitivity. However, as most other studies, they exclude (see their Table 2, footnote b) the amounts spent on cinema and outdoor media. As their share tends to be smaller, very few studies have explicitly considered the effectiveness of billboard and/or cinema advertising. Two notable exceptions are Berkowitz, Allaway and D'Souza (2001) and Naik, Peters and Raman (2008). The former analyzed one year of weekly sales, radio advertising and billboard advertising for three stores of a single regional retailer, and concluded (without reporting specific elasticities) that "radio advertising is anywhere from three to seven times (depending on the store) more effective than billboard advertising" (p. 64). Naik et al. (2008) considered a multimedia campaign for soft drinks, and reported a greater impact of TV advertising (elasticity of 0.32) relative to print (0.02), outdoor (0.06) and cinema (0.06). Again, it is hard to interpret this lower effectiveness of billboard and/or cinema advertising as an empirical regularity on the basis of just two studies, especially since the reported elasticity of TV advertising in the Naik et al. (2008) application appears to be much higher than the average value reported in the aforementioned meta-analyses, which could be attributed to their less conventional dependent variable (i.e., medium-specific advertising awareness levels rather than sales or market share).

Against this background, our study offers new substantive and managerial insights. *Substantively*, we contribute novel empirical generalizations on the effectiveness of advertising in two media that have been largely ignored in prior research by systematically analyzing the short- and long-run advertising effectiveness of both billboard spending and cinema advertising for a wide variety of CPG brands (40+ for cinema advertising; 100+ for billboard advertising). While the proportion spent worldwide on those media (estimated at 6.6% for outdoor and 0.6% for cinema in 2012; Barnard 2012) may be smaller than for the more often studied television (40.2%) and print (27.7% for newspapers and magazines combined) media, the absolute spending levels remain very large (32.3 US$ billion and 2.7 US$ billion, respectively), and warrant more research attention. For the more traditional media, this paper is the first to provide empirical generalizations on sales, as well as market-share, elasticities that have been corrected for the brands' self-selection in media usage. *Managerially*, a better understanding of their relative effectiveness will help managers make better media-allocation decisions, not only managers who currently abstain from using those media, but also those who already allocate a substantial portion of their media budget to them.

2. DATA

Through GfK Benelux (monthly revenue) and the Belgian Centre for Information about Media (advertising expenditures across media), we obtained monthly sales and advertising-spending information on 261 leading brands in the consumer packaged goods (CPG) market. The brands cover a wide variety of 96 categories, involving food, beverage, household-care, personal-care and pet-food products (a summary is provided in Table 1). The revenue data consist of the aggregated data across all SKUs of the brand within the category at hand (e.g., combined sales across all Axe deodorants). Special care was taken that the advertising data were compiled according to the same category classification. All brands advertised at least two times during the observation period (January 2004 – August 2010) in at least one of the following six advertising media: television, radio, newspapers, magazines, billboards and cinema.[1] Sales revenues and advertising expenditures were deflated based on the relative Consumer Price Index (CPI).

--- Table 1 about here ---

As shown in Tables 2 and 3, there is considerable variability in media usage. Table 2 illustrates how the different media differ in both (i) their number of users[2] and (ii) their share of the total advertising budget (adspend share). In line with previous research (see, e.g., Deleersnyder et al. 2009; Sethuraman et al. 2011), TV advertising is by far the most popular medium: 231 out of the 261 brands use TV advertising, resulting in a combined spending share of 82.9%. All other media are considerably less popular. Billboard (cinema) advertising, for example, has an adspend share of only 7.2 (1.8)%, and is used by only 117 (43) brands. Among those users, the

[1] No internet spending was considered. The data collection on this medium was started two years after that on the other media, which would have reduced considerably the length of the time series.
[2] A brand is considered a user of a certain advertising medium if it uses the medium at least two times during the observation period (a similar decision rule was recently used in van Heerde et al. 2013).

corresponding adspend shares are considerably higher, however (column 4 of Table 2), and even exceed (with a share of 11.9 and 6.9%) the proportion spent on more traditional media such as radio (5.7%), newspapers (3.1%) and magazines (4.4%). Interestingly, brands differ considerably in the number of media they use, as summarized in Table 3. While 52 (20%) brands use only a single medium, other brands (18) use all six media.

--- Tables 2 and 3 about here ---

Besides sales revenues,[3] GfK provided information on the brands' penetration level and market share. These variables will be used to address potential sample-selection issues (we refer to the Methodology Section for details), and to explore differences in both media usage and media effectiveness.

3. METHODOLOGY

Our modeling approach consists of two steps: (i) an analysis at the brand level to estimate the advertising elasticities for the different media used by each individual brand, and (ii) a meta-analysis to combine these results into empirical generalizations. In our first step, i.e. the brand-specific analysis, we closely follow van Heerde et al. (2013), and use an error-correction model to estimate each medium's short- and long-run advertising elasticity, while we correct for endogeneity (and allow for correlated error terms within a given product category) through a 3SLS estimation procedure. However, unlike van Heerde et al. (2013), where *all* brands made

[3] No price information was available. However, cross-sectional differences in average price level across brands in a given category are captured through our brand-specific intercepts, as explained in the Methodology section. Moreover, Sethuraman et al. (2011, Table 2) show in their meta-analysis on brand-level advertising elasticities that the absence of a price variable in the model has no significant impact on the resulting advertising-effectiveness estimates, neither in the short run nor in the long run.

use of the (aggregated across media) advertising instrument, we are confronted with an intrinsic selection problem. Indeed, many brands make no use of certain advertising channels (as documented in Table 3). We explicitly account for this selection issue in our second-stage meta-analysis.

3.1 Brand-specific analysis

An error-correction model is used to estimate both the short- and long-run advertising elasticities of each medium used by each individual brand. If *i* represents the brand (i.e., $i = 1,\ldots,B_c$) and *c* the category the brand belongs to ($c = 1,\ldots,96$), the error-correction model is written as:

$$\Delta \ln \mathrm{Revenue}_t^{ic} = \alpha^{ic} + \sum_{j=1}^{J^{ic}} \beta_j^{ic} \Delta \ln \mathrm{Advertising}_t^{icj} + \beta_{J^{ic}+1}^{ic} \Delta \ln \mathrm{Other}_t^{ic}$$

$$+ \varphi^{ic} \left[\ln \mathrm{Revenue}_{t-1}^{ic} - \textstyle\sum_{j=1}^{J^{ic}} \gamma_j^{ic} \ln \mathrm{Advertising}_{t-1}^{icj} - \gamma_{J^{ic}+1}^{ic} \ln \mathrm{Other}_{t-1}^{ic} - \delta^{ic} \mathrm{trend} \right] + \varepsilon_t^{ic}, \quad (1)$$

where J^{ic} denotes the number of media for which brand *i* from category *c* had (in line with the decision rule of van Heerde et al. 2013) at least two non-zero spending levels. The β_j^{ic} coefficients represent the short-run (same period) advertising-to-sales elasticities of TV, radio, newspaper, magazine, billboard or cinema, as $\Delta \ln \mathrm{Advertising}_t^{icj}$ gives the first difference of the log-transformed advertising expenditures. The γ_j^{ic} parameters represent their long-run counterparts, and reflect the cumulative effect (same period + future periods) of a one-period shock to the medium at hand.[4] $\beta_{J^{ic}+1}^{ic}$ and $\gamma_{J^{ic}+1}^{ic}$ capture the combined short-run and long-run advertising elasticity across media that were used only once in the observation period. The

[4] This error-correction interpretation presumes stationarity of the (log) sales series, which was confirmed through the Levin, Lin and Chu (2002) panel unit-root test. A similar conclusion was obtained on the basis of brand-specific Phillips-Perron (1988) tests (detailed results are available from the first author upon request).

φ^{ic} parameter reflects the speed of adjustment towards the underlying long-run equilibrium, and the trend variable serves as a proxy for all other variables that have gradually changed over the sample period (cf. Dekimpe and Hanssens 1995). This trend variable ranges from -1 to +1 in order to interpret the elasticities mid-observation period. An in-depth discussion of the error-correcting specification is provided in Fok et al. (2006). Its use is well-established in the marketing literature (see, e.g., Horváth and Fok 2013; van Heerde et al. 2007, 2013 for other applications), and is particularly suited for our research setting, as it provides direct estimates of the different media's short- and long-run elasticities, without the need to impose a common carryover coefficient (as is often done in Koyck-type and partial-adjustment-type models involving multiple media; see Naert and Leeflang 1978, pp. 94-97).[5] 3SLS estimation is used to account for the possible endogeneity of the $\Delta \ln \mathrm{Advertising}_t^{icj}$ variables.[6] The averaged values of the lagged log advertising expenditures from the other product classes, disaggregated across the six advertising media, serve as instruments (see van Heerde et al. 2013 or Lamey et al. 2012 for a similar practice). As we have 24 instruments (six media across four non-focal product classes) for a maximum of six endogenous variables, the model is overidentified. The model is estimated jointly across the B_c brands within a given category c, allowing their error terms to be correlated.

3.2 Meta-analysis

In the second step of our analysis, we meta-analytically combine the brand-specific parameter estimates, and derive empirical generalizations on the relative effectiveness of the different

[5] Note that the autoregressive partial-adjustment model will be applied as a robustness check in Section 4.4.
[6] $\ln \mathrm{Advertising}_{t-1}^{icj}$ and $\ln \mathrm{Other}_{t-1}^{ic}$ are predetermined variables, and are thus not endogenous. We treat $\Delta \ln \mathrm{Other}_t^{ic}$ as exogenous, as too few non-zero observations might prevent a reliable auxiliary-regression estimation (Steenkamp et al. 2005).

media. Unlike van Heerde et al. (2013), we cannot rely on Rosenthal's method of added Zs (Rosenthal 1991). As we estimate the elasticities at a more disaggregate level (i.e., at the level of the individual medium, rather than summed across all media --- in which case all included brands are a user), and given that not all media are used by every brand, there is a potential issue of sample selection. As brands self-select which media to use, only those brands for which a medium is likely to be (more) effective may do so, and their estimates may not generalize to the other brands. As such, we have to test (and correct) for this potential self-selection bias. In addition, we need to account for the fact that (i) the elasticities are estimated quantities, and (ii) that estimates from brands within the same category may not be independent of one another.

Two meta-regressions are constructed for, respectively, the stacked short-run elasticity estimates (2a) and the stacked long-run estimates (2b) from the brand-specific analyses:

$$\beta_{m}^{ic} = \sum_{k=1}^{6} \beta_{k}\, D_{k}^{ic}(m) + \sum_{k=1}^{6} \rho_{k}^{SR}\, \lambda_{k}^{ic}\, D_{k}^{ic}(m) + u_{m}^{ic}, \tag{2a}$$

$$\gamma_{m}^{ic} = \sum_{k=1}^{6} \gamma_{k}\, D_{k}^{ic}(m) + \sum_{k=1}^{6} \rho_{k}^{LR}\, \lambda_{k}^{ic}\, D_{k}^{ic}(m) + v_{m}^{ic}. \tag{2b}$$

The index *m* denotes the medium at hand (TV, radio, newspapers, magazines, billboards or cinema), with each brand contributing J^{ic} elasticity estimates. $D_{k}^{ic}(m)$ is an indicator variable, taking on the value 1 when $k = m$. As such, β_{k} and γ_{k} provide the expected (short- and long-run) effectiveness of each medium *k*.

However, their estimates may not represent the expected short- and long-run advertising elasticities for a randomly-selected CPG brand because of the aforementioned self-selection into the sample. In fact, they would only pertain to those brands that actually use the medium. To correct our meta-analytic results for this potential selection bias, we include in Equations (2a) and (2b) λ_k^{ic} variables resulting from Heckman's two-stage method (Greene 2000). For each of the six media, a probit model is estimated to quantify the probability that the medium is used by brands with certain characteristics:

$$\Pr\left(s_m^{ic} = 1 \,\middle|\, z_m^{ic}\right) = \Phi(z_m^{ic}\,\xi_m). \tag{3}$$

The dummy variable s_m^{ic} is a selection indicator, which takes on the value 1 when brand *i* from CPG category *c* is a user of medium *m*. z_m^{ic} is a 1 x 12 vector of variables that might have explanatory power over the decision whether or not a brand is a user of a given medium. This vector includes (i) an intercept, (ii) four product-class dummies indicating whether or not the brand is a beverage, food, personal-care or household-care brand (with pet food as reference category), (iii) the average values of the log advertising expenditures for each of the five other media, (iv) the average market share of the brand during the observation period, and (v) its average penetration level. Using the fitted values of this probit model, the corresponding inverse Mills ratio is derived as:

$$\hat{\lambda}_m^{ic} = \frac{\phi(z_m^{ic}\hat{\xi}_m)}{\Phi(z_m^{ic}\hat{\xi}_m)}, \tag{4}$$

which is added to the meta-regression (2) as an additional explanatory variable.

When estimating (2), one needs to account for the fact that the dependent variables β_m^{ic} and γ_m^{ic} are estimated quantities, as are the inverse Mills ratios. Moreover, the error terms corresponding to brands from the same category may not be independent of one another. To accommodate the first issue (estimated dependent variables), we use weighted least squares (WLS) to obtain parameter estimates in the meta-regressions. The weights are set to equal the inverse of the standard errors of the β_j^{ic} and γ_j^{ic} estimates from the 3SLS estimation (see

Saxonhouse (1976) for a formal motivation, or Nijs, Srinivasan and Pauwels (2007) for another marketing application). To address the two other issues (estimated explanatory variables and correlated error terms), the standard errors for the meta-analytic parameters are derived by means of a bootstrap procedure, where we also account for clustered error terms between brands of the same product category (Field and Welsh 2007).

4. RESULTS

4.1 Small-medium usage

All six probit selection models (3) resulted in a high hit rate (ranging between 71 and 89%), which was substantially higher than the number expected by chance (Morrison 1969). Even though not our primary research interest, we obtained several interesting insights (detailed results are provided in Online Appendix A) into which brands/categories are most likely to *use* the less popular billboard and cinema media. First, brands with a higher penetration level (two-sided $p <$ 0.10) are more likely to use the billboard medium.[7] In contrast, high-share brands are not more/less likely ($p > 0.10$) to use billboard or cinema advertising than smaller brands. Interestingly, brands that make more extensive use of the TV medium are also more likely (two-sided $p < 0.05$) to use cinema advertising, while brands that make more extensive use of the magazine medium are also more likely (two-sided $p < 0.05$) to use billboard advertising. These combinations make intuitive sense, as tactical consistency (Sheehan and Doherty 2001) may be easier to realize between TV and cinema advertising on the one hand, and between magazine and billboard advertising on the other hand, which should facilitate consumers' information processing when these media pairs are used together in a coordinated media campaign (Edell and

[7] This is also the case for television advertising.

Keller 1989). Finally, significant differences between the different category types are present as well. Billboard advertising, for example, is more likely to be used by brands in the food and beverage category, and less likely in the household-care, personal-care and pet-food category, while beverage brands are typically more inclined to use cinema advertising compared to food brands.

4.2 Brand-specific findings

Table 4 summarizes the individual-level elasticity estimates. For seven brands, preliminary single-equation estimations of the sales equation resulted in a variance inflation factor (VIF) in excess of 10, indicating serious multicollinearity issues (Hair et al. 1995). We omit these brands when reporting the proportion of significant effects (and from our subsequent meta-analysis).[8] Focusing on the long-run parameter estimates for the remaining 254 brands, we find that 141 (18.7%) of the 755 estimated advertising elasticities are positive and significant ($p < 0.10$, one-sided)[9]. This proportion is comparable to the one obtained in other CPG-based studies (see, e.g., van Heerde et al. 2013 for the U.K., and Steenkamp et al. 2005 for the Netherlands). Interestingly, the proportion of significant estimates for TV (28%) is significantly higher ($p < 0.05$) than for the other media, offering a first justification for its more frequent use. Still, even for that medium, no significant effect is found for a substantial number of brands, in line with earlier findings (see, e.g., Lodish et al. 1995; Allenby and Hanssens 2005) that advertising for mature products often fails to induce significant sales increases. Moreover, a large number of brands (143 or 56.3%) does not obtain a significant effect with any medium, and only 26 brands (out of the 202 that use at least two media) have a significant long-run elasticity for more than

[8] We do so because the resulting inflation of the standard errors would render the estimation results too imprecise.
[9] In Online Appendix B, we also report the proportion of positive (not necessarily significant) estimates.

one (with a maximum of three) medium. This again suggests that, at least from a sales-response perspective, advertising spending can often be perceived, in the terminology of Leeflang and Wittink (1996) and Steenkamp et al. (2005), as spoiled arms.

--- Table 4 about here ---

Considering our two focal media, billboards (12.9%) and cinema (15.0%), their proportion of significant effects is not significantly lower ($p > 0.10$) than for the more traditional radio (18.4%), newspaper (12.8%) and magazine (15.5%) media. These numbers exceed (albeit only marginally) the percentage that would be expected by chance (10%), but are (as indicated before), significantly smaller than the proportion obtained for TV advertising.

Only 26 of the 202 brands that actively use more than one medium obtain a significant long-run elasticity for multiple media. This raises serious concerns about the optimality of the allocation rules used by many brands. Indeed, almost all brands allocate resources to one or more non-effective (from a sales-response point of view) media. For only 11 of these 26 brands, one of the significant long-run elasticities involved billboard (8) and/or cinema advertising (4). Comparing the relative share allocated to these media (e.g., spending on billboard relative to the spending on all "effective" media) with the near-optimal allocation heuristic developed in Fisher et al. (2011),[10] we note (i) that among the eight billboard users, five brands are "on target". They allocate a relative adspend share to billboard advertising (relative referring to the shares allocated to the other significant media) that lies within the 90% confidence interval for the near-optimum. The other brands significantly underspend on this type of advertising, with an average deficit in relative budget share of 39.6 percentage points. For cinema advertising, the results on adspend

[10] According to Fisher et al.'s heuristic, this proportion should be equal to the ratio between that medium's own long-run elasticity and the sum of all the significant long-run elasticities.

share allocation are not as conclusive (one brand is on target, while one (two) brand(s) significantly overspends (underspend) on cinema advertising).

4.3 Meta-analytic findings

One could argue that the rather bleak picture on advertising's effectiveness may be attributed to the low power of the statistical tests, as "only" 80 observations are available for each brand. To increase the power of our inference, we meta-analytically combine the evidence across all brands that use a particular medium (see, e.g., Deleersnyder et al. 2004, 2009; Lamey et al. 2012; van Heerde et al. 2013 for a similar reasoning). In both the short- (television, radio and newspapers) and the long-run (television, newspapers, billboards) equation, several of the ρ_k estimates are significant ($p < 0.10$, one-sided), indicating the need to control for self-selection. This was also confirmed when looking at the combined evidence across the six ρ parameters by means of the Strube (1985) test (see Deleersnyder et al. 2004 for a marketing application), which was highly significant in both instances (one-sided $p = 0.005$ and 0.013 for Equations (2a) and (2b), respectively).

The resulting elasticity estimates are summarized in Table 5. In both the short and long run, the overall advertising elasticity is only significant for TV and magazines. In other words, for a random CPG brand, TV and magazines are the only media that are able to significantly affect the brands' sales revenues, with television the most effective ($p < 0.001$) of the two. For the other media, including billboards and cinema, no meta-analytic evidence of their effectiveness is found (short-run p-values of 0.152 (billboards) and 0.317 (cinema); long-run p-values of 0.451 and 0.148, respectively).

Comparing the meta-analytic results *with* and *without* correction for self-selection, we find that in the latter case, the estimate for the long-run TV elasticity is biased upwards (0.0070 versus 0.0043). Moreover, for the other media, a significant long-run effectiveness is also found for radio advertising and billboard advertising (in terms of short-run effectiveness, three additional effects become significant, i.e. radio, newspapers and billboards). Hence, not accounting for the fact that the observed set of users of a given medium may not be fully random leads one to seriously overestimate its effectiveness when generalizing to the overall population of brands.

--- Table 5 about here ---

4.4 Robustness checks

Alternative performance metric: market share. The earlier discussion focused on the impact of the different media on the brands' sales revenues, which can reflect both competitive gains and gains because of primary market expansion (Sethuraman et al. 2011). In a first robustness check, we replaced the dependent variable $\Delta \ln \mathrm{Revenue}_t^{ic}$ in the brand-specific equation (1) by $\Delta \ln \mathrm{MarketShare}_t^{ic}$, which only captures the competitive gains. All other steps of the analytic approach remained unchanged,[11] except that we now used the brand's revenues (rather than its share) as an explanatory variable in the probit selection models for media usage. The results were very robust across both dependent measures: (i) no significant effect was found for the two focal small media, i.e. billboard and cinema, neither in the short run (Table 6a) nor in the long

[11] As Leeflang and Wittink (1996), we did not consider the logical-consistency requirements in our market-share validation, for two reasons. First, empirical evidence on the relative performance of multiplicative versus attraction specifications remains inconclusive (we refer to their footnote 10 for an extensive literature review on the issue). Second, for several categories, information on only a small subset of the brands in the category was available, making a full attraction specification infeasible. Also, given that our data set contains a large number (96) of micro-categories, the approach used in Fok, Paap and Franses (2003), where one would use one brand per (micro-) category as reference brand, would lead to a substantial reduction in the number of elasticities that could be obtained per medium, making also that approach less appealing.

run (Table 6b), (ii) this was also the case for radio and newspaper advertising, while (iii) a significant effect was again found for TV advertising. The only exception was the magazine medium, where the significant effect in terms of sales revenue appears to be attributable to a primary-demand, rather than a competitive-gain, effect (as the impact becomes insignificant in the market-share model).

--- Table 6 about here ---

Alternative model specification: autoregressive partial-adjustment model. To ensure that the lack of empirical evidence on the short-run and long-run effectiveness of the small advertising media is not idiosyncratic to the error-correction specification used in Equation (1), we re-estimated the brand-specific advertising elasticities using the autoregressive partial-adjustment model (APAM) discussed in Hanssens, Parsons and Schultz (2001, p. 147) and Leeflang et al. (2000, p. 489). The APAM specification is very similar to the Koyck model, but is somewhat more flexible, in that it does not impose the restriction that the autocorrelation parameter in the error term equals the coefficient associated with the lagged dependent variable. Specifically:

$$\ln \mathrm{Revenue}_t^{ic} = \left(1-\lambda^{ic}\right)\left[\alpha^{ic} + \sum_{j=1}^{J^{ic}} \beta_j^{ic} \ln \mathrm{Advertising}_t^{icj} + \beta_{J^{ic}+1}^{ic} \ln \mathrm{Other}_t^{ic}\right]$$

$$+ \lambda^{ic} \ln \mathrm{Revenue}_{t-1}^{ic} + \varepsilon_t^{ic}, \qquad (5)$$

with $\varepsilon_t^{ic} = \tau^{ic}\, \varepsilon_{t-1}^{ic} + w_t^{ic}$, denoting that w_t^{ic} and ε_t^{ic} follow a white-noise respectively AR(1) process. In this specification (see also Bass and Clarke 1972), the β_j^{ic} coefficients represent the long-run advertising-to-sales elasticities of TV, radio, newspaper, magazine, billboard or cinema, while the short-run elasticities can be calculated as $\left(1-\lambda^{ic}\right)\beta_j^{ic}$. λ^{ic} gives the brand-specific carryover parameter that is instrumental in linking the short- and long-run elasticities. Again, similar conclusions were obtained for the two focal small media (with no significant meta-analytic effect in the short or long run), for radio and magazine advertising (with no significant effect either), and for television (with a significant short- and long-run effect ($p < 0.01$)). For magazine advertising, a comparable short- (0.0024 versus 0.0019) and long-run (0.0041 versus 0.0059) elasticity as before was obtained, even though the latter failed to reach statistical significance in the APAM specification.

Alternative diminishing returns to scale. The error-correction model was specified in the popular log-log format, which offers direct estimates of the respective elasticities, and which captures diminishing returns to scale. To verify that our main insights remain unchanged with a different rate of diminishing returns to scale, we adopted the specification of Naik and Raman (2003), where we model the dependent variable (sales revenues) as a function of the square root of the media expenditures (see also Leeflang et al. 2000, p. 69). Across all six media, the same conclusions[12] as in our focal model were obtained, i.e. insignificant meta-analytic effects for billboard, cinema, newspaper and radio, and a significant short- and long-run effect for TV and magazines. This was also the case when we varied, in the spirit of Lamey et al. (2007, Table 4), the rate of diminishing returns further, and used $\text{Advertising}^{0.4}$ and $\text{Advertising}^{0.6}$ instead of $\text{Advertising}^{0.5}$.

Stability. As a final robustness check, we tested the stability of the obtained insights on media effectiveness through a split-half analysis. To that extent, Equation (1) was augmented as follows:

[12] Given the different scaling of both the dependent (log-transformed or not) and independent (log-transformed or square root) variables, the parameter estimates are not directly comparable in magnitude across both specifications. As such, we focus on the sign and significance in this comparison.

$$\Delta \ln \text{Revenue}_t^{ic} = \alpha^{ic} + \sum_{j_1=1}^{J_1^{ic}} \beta_{1,j}^{ic}\, \Delta \ln \text{Advertising}_t^{icj}\, S_t^1 + \sum_{j_2=1}^{J_2^{ic}} \beta_{2,j}^{ic}\, \Delta \ln \text{Advertising}_t^{icj}\, S_t^2$$

$$+ \beta_{1,J_1^{ic}+1}^{ic} \Delta \ln \text{Other}_t^{ic}\, S_t^1 + \beta_{2,J_2^{ic}+1}^{ic} \Delta \ln \text{Other}_t^{ic}\, S_t^2$$

$$+ \varphi^{ic} \Big[\ln \text{Revenue}_{t-1}^{ic} - \textstyle\sum_{j_1=1}^{J_1^{ic}} \gamma_{1,j}^{ic} \ln \text{Advertising}_{t-1}^{icj}\, S_t^1 - \sum_{j_2=1}^{J_2^{ic}} \gamma_{2,j}^{ic} \ln \text{Advertising}_{t-1}^{icj}\, S_t^2 -$$

$$\gamma_{1,J_1^{ic}+1}^{ic} \ln \text{Other}_{t-1}^{ic}\, S_t^1 - \gamma_{2,J_2^{ic}+1}^{ic} \ln \text{Other}_{t-1}^{ic}\, S_t^2 - \delta^{ic} \text{trend} \Big] + \varepsilon_t^{ic}, \tag{6}$$

with S_t^1 and S_t^2 dummy variables representing the two subsamples.[13] J_1^{ic} and J_2^{ic} denote the number of media for which brand *i* from category *c* had at least two non-zero spending levels in, respectively, subsample 1 and 2. Subsequently, we estimated subsample-specific (sample-selection-corrected) weighted averages of the brand-specific elasticities per medium:

$$\beta_{m,s}^{ic} = \textstyle\sum_{l=1}^{2} \sum_{k=1}^{6} \beta_{k,l}\, D_k^{ic}(m)\, D_l^{ic}(s) + \sum_{k=1}^{6} \rho_k^{SR}\, \lambda_k^{ic}\, D_k^{ic}(m) + u_m^{ic}, \tag{7a}$$

$$\gamma_{m,s}^{ic} = \textstyle\sum_{l=1}^{2} \sum_{k=1}^{6} \gamma_{k,l}\, D_k^{ic}(m)\, D_l^{ic}(s) + \sum_{k=1}^{6} \rho_k^{LR}\, \lambda_k^{ic}\, D_k^{ic}(m) + v_m^{ic}, \tag{7b}$$

where $\beta_{m,s}^{ic}$ ($\gamma_{m,s}^{ic}$) consist of the estimated short-run (long-run) elasticities of the different media for both subsamples (i.e., *s* = 1, 2 refers to the first, respectively second, subsample). $D_l^{ic}(s)$ is an indicator variable, taking on the value 1 when *l* = *s*. The final columns in Table 6 report the respective estimates. As before, similar substantive conclusions were obtained, with insignificant effects (both in the short and in the long run) for billboard, cinema, radio and newspaper advertising (in both subsamples), and a highly significant ($p < 0.01$) effect for TV advertising (again in both subsamples). In addition, three of the four magazine elasticities were found to be significant ($p < 0.10$). Moreover, none of the elasticities found in the second half of

[13] $S_t^1 = 1$ when $1 \le t \le 40$ while $S_t^2 = 1$ when $40 < t \le 80$.

the sample was significantly different ($p > 0.10$) from the corresponding value in the first subsample, which again confirms the stability of our inferences over time.

In sum, our results were found to be very robust across all sensitivity analyses. We consistently found no significant meta-analytic effect for the two focal small media, nor for the more traditional radio and newspaper media. In contrast, we each time found a significant short- and long-run effect for TV advertising. The impact of magazine spending, in turn, while robust in sign and magnitude, sometimes failed to reach statistical significance.

4.5 Heterogeneity across brands and categories

In the spirit of van Heerde et al. (2013), we also derived the small media's (average) effectiveness in some major product groups (beverages, food, household care, personal care, and pet food for billboards, and beverages, food and personal care for cinema).[14] No significant meta-analytic elasticity was found (all p-values > 0.10) in any of these category types, neither in the short nor in the long run. We also checked (one at a time)[15] for a moderating impact of the brands' (i) average market share, (ii) penetration level, and (iii) advertising intensity (measured as the combined spending across all six media relative to the brands' sales revenue). However, the total (meta-analytic) impact of the two small media (main + moderating effect) never reached statistical significance (i.e., $p > 0.10$) across the range of these moderators observed among their users.

---

[14] No pet-food or household-care brand in our sample made use of cinema advertising.
[15] Specifically, we added interaction terms between a given moderator and the six media dummies to Equation (2).

4.6 How about synergy effects?

Even in the absence of significant "own" effects, allocating a portion of one's advertising budget to a certain medium may be justified when it enhances the effectiveness of another medium (Naik and Raman 2003; Naik 2007, p. 44). To this end, we investigated to what extent billboard and cinema advertising have a positive synergistic effect with television, the most frequently used (Table 2) and most effective (Tables 4 and 5) medium.

Specifically, we augmented Equation (1) with an interaction term between both media, both in the short-run (i.e., between the first difference terms) and the long-run (i.e., between the lagged level terms) part. 70 brands were a joint user of both TV and billboard (i.e., had at least two joint spending occurrences), of which 41 had a VIF value smaller than 10 after the inclusion of the relevant interaction terms. Of these 41 brands, 3 (6) experienced a significantly positive ($p <$ 0.10, one-sided) short-run (long-run) synergy effect from this joint spending, a proportion not exceeding what could be expected by chance (to correct for potential self-selection, we estimated, in a similar spirit as before, a probit model on an indicator variable taking the value of 1 if the brand was a user of both media, and zero otherwise, to derive the corresponding inverse Mills ratio). Also meta-analytically, no significant synergy effect was found in the short or long run ($p > 0.10$). As for cinema advertising, only 16 brands passed both criteria (at least two joint occurrences, and no evidence of serious multicollinearity). In this case, only 2 (0) brands experience a significant long-run (short-run) synergy effect. Across the 16 brands, no significant meta-analytic effect was found ($p > 0.10$). Hence, little evidence is found in support of the synergy claim that is often raised (see, e.g., Bhargava and Donthu 1999, p. 8; Ewing, du Plessis and Foster 2001, p. 78) to motivate a more extensive use of these smaller media.

Following a similar procedure, we checked (one media combination at a time)[16] for a positive synergy effect between our two focal media (billboard and cinema) and the three other media (radio, newspapers, and magazines). Table 7 summarizes the various synergy tests. For cinema advertising, no significant synergy effect was obtained (neither in the short nor in the long run) with any of these media ($p > 0.10$ in all instances). This was also the case for billboard advertising when combined with newspaper and magazine advertising. However, a positive long-run synergy effect, estimated at 0.005, was found between billboard and radio advertising. This effect was estimated on 24 brands that used both media simultaneously in at least two instances (and for which the VIF values stayed below 10 after the inclusion of the relevant interaction terms). The synergy effect was significant ($p < 0.10$) for four individual brands, as well as meta-analytically ($p < 0.05$). Even though the long-run main effect stayed insignificant,[17] the *total* impact of billboard spending (consisting of the main and moderated impact) was significant ($p < 0.10$) in case of positive spending levels on radio. Still, brands often seem to ignore this opportunity. Indeed, many brands (89) used billboard advertising without ever using it simultaneously with radio spending. Across the 24 brands that did use both media jointly, billboard advertising coincided with radio spending in only 27% of the time periods that the billboard medium was used (and in only 4% of all possible calendar months). Hence, in many instances, brands fail to capitalize on the potential synergy effects that a joint use of both media could entail.

---

[16] Retaining all possible interaction terms jointly would lead to excessive multicollinearity and unstable results. In the spirit of Bijmolt, van Heerde and Pieters (2005), we opted instead to add one media combination at a time to our base model (1). We refer to van Heerde et al. 2013, Steenkamp and Geyskens 2014, or Yeung et al. 2013 for a similar practice.

[17] The short-run main effect of billboard, in contrast, became significant ($p < 0.05$) after allowing for an interaction effect for those 24 brands.

Finally, we also checked for a synergy effect between both small media, but no such evidence ($p > 0.10$) was found, neither in the short run nor in the long run.

--- Table 7 about here ---

5. CONCLUSION

Even though numerous studies have quantified the sales effectiveness of advertising spending, most have either focused on *aggregated* spending (i.e., summed across media), or on the most popular (*television*) medium (e.g., Sethuraman et al. 2011). In this study, we provide some new empirical generalizations on the relative effectiveness of less-frequently used, and definitely much less studied, media such as billboard and cinema advertising. Importantly, we show the need to correct for self-selection when making inferences, not only on the effectiveness of these smaller media, but also on the more traditional media.

Using a rich data set on over 250 popular CPG brands, we find very little evidence in support of the small media's sales effectiveness: (i) a significant short- and/or long-run elasticity is found for only a small fraction (16.4 and 12.9% for billboards; 7.5 and 15.0% for cinema advertising) of brands, and (ii) also meta-analytically, no significant positive effect ($p > 0.10$) was obtained. Moreover, no evidence of synergy effects with the most popular and most effective medium (TV) was found --- precluding another justification for their use. Only with one other medium (radio) did we find evidence of a positive (long-run) synergy effect. However, only in very few instances (both in terms of the number of brands and in terms of the number of time periods) did brands try to capitalize on this opportunity.

These results may sound discouraging. Indeed, both managers and academics prefer significant results with large effect sizes. However, as emphasized in Hubbard and Armstrong (1992) and

Sawyer and Peter (1983), there is clear value in the identification of a null result, especially when this result is obtained in many replications (in our case, across many brands and categories), and when this also holds up in the much more powerful meta-analytic tests (see also Fagley 1985). Moreover, it is important to realize that our results (both in terms of the proportion of significant effects and in terms of the average effect size) are in line with many of the earlier findings on advertising's limited sales effectiveness in mature CPG categories (e.g., van Heerde et al. 2013; Srinivasan et al. 2010). This could be interpreted as evidence that advertising spending in general, and in the smaller billboard and cinema media in particular, reflects ineffective/spoiled arms in those markets.[18] However, it is interesting to note that, without self-selection correction, the meta-analytic effect for billboards was significant (cf. Table 5). Hence, the (small) subset of brands that makes use of this medium appears well-informed (on average) that this medium works for them.[19] Also, one should not forget that advertising (also in those smaller media) may well have other benefits, such as a reduced vulnerability to competitive actions (van Heerde et al. 2007), a lower sensitivity to cyclical fluctuations (Deleersnyder et al. 2009), a lower private-label growth (Lamey et al. 2012), an enhanced visibility to financial stakeholders (Joshi and Hanssens 2010), or the stimulation of word-of-mouth communication among potential customers, which is an essential element in any successful grassroots marketing campaign (Stephen and Galak 2012), to name just a few (see also Allenby and Hanssens 2005 or Dekimpe and Hanssens 2011 for a similar reasoning). Also, we focused on the sales (and market-share) effectiveness of the different media. It would be interesting to also study, in the spirit of Srinivasan et al. (2010) and

[18] Strictly speaking, we were not able (in spite of many replications and robustness checks) to prove the null hypothesis of no sales impact wrong, which does *not* imply, however, that the null hypothesis has been proven true. Still, we follow common practice (see, for example, Leeflang and Wittink 1996; Steenkamp et al. 2005) of calling a marketing instrument ineffective (or spoiled arms) if no significant sales elasticity is found.
[19] We would like to thank the Associate Editor for this observation.

Pauwels, Erguncu and Gokhan (2013), how they influence intermediate mind-set metrics, such as brand recognition or brand liking.

Also within the sales-responsiveness domain, multiple avenues for future research remain. First, it would be useful to control for other marketing-mix instruments, such as distribution and salesforce, when estimating the brand-specific advertising elasticities in Equation (1). However, these are likely to be positively correlated with current-period sales and with current-period advertising. As such, one can expect the omission of distribution (salesforce) *to bias the advertising-elasticity estimate positively* (Sethuraman et al. 2011, p. 461, italics added). This was indeed found in the meta-analysis of Sethuraman et al. (p. 462, p. 468). Given this result, the absence of any significant effect in our model without these two measures makes our "spoiled arms" conclusion a conservative one.[20] Still, we agree that it would be useful to add specific (also longitudinal) information on those variables. Second, and in spite of their lower overall effectiveness, significant billboard or cinema effects are obtained for certain brands. Indeed, 20.7% (17.5%) of the users obtained either a significant short- or long-run response from their investments in billboard (cinema) advertising. In several instances, this effect was even larger than for any of the other media used by the brand. Hence, more research is warranted not only to identify (beyond the potential moderators studied in Section 4.5) for what brands, categories and/or market conditions the occurrence of such positive effects becomes more likely, but also to identify what creative (more qualitative) aspects of the campaign make a significant main and/or synergy effect for a particular medium more pronounced. Third, we focused on

[20] Also, in line with our earlier footnote 3, cross-sectional differences in average distribution/salesforce support across brands in a given category are already reflected in the brand-specific intercepts.

established brands in mature CPG categories. More research is needed whether stronger effects for billboard and/or cinema advertising are found for newer brands, or with non-CPG categories.

REFERENCES

AdAge (2012). Advertising to sales ratios by industry. Posted June 28, 2012. Available at http://adage.com/article/datacenter-advertising-spending/advertising-sales-ratios-industry/106575/ .

Allenby, Greg, & Hanssens, Dominique M. (2005). Advertising response. *Marketing Science Institute, Special Report*, (05-200), 8.

Barnard, Jonathan (2012). ZenithOptimedia forecasts 4.1% growth in global adspend in 2013. Posted December 3, 2012. Available at http://www.zenithoptimedia.com/zenith/zenithoptimedia-forecasts-4-1-growth-in-global-adspend-in-2013/.

Bass, Frank M., & Clarke, Darral G. (1972). Testing distributed lag models of advertising effect. *Journal of Marketing Research*, *9*(3), 298-308.

Berkowitz, David, Allaway, Arthur, & D'Souza, Giles (2001). The impact of differential lag effects on the allocation of advertising budgets across media. *Journal of Advertising Research*, *41*(2), 27-38.

Bhargava, Mukesh, & Donthu, Naveen (1999). Sales response to outdoor advertising. *Canadian Journal of Marketing Research*, *12*(1), 87-97.

Bijmolt, Tammo H.A., Van Heerde, Harald J., & Pieters, Rik G.M. (2005). New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, *42*(2), 141-156.

Danaher, Peter J., & Dagger, Tracey S. (2013). Comparing the relative effectiveness of advertising channels: a case study of a multimedia blitz campaign. *Journal of Marketing Research*, *50*(4), 517-534.

Dekimpe, Marnik G., & Hanssens, Dominique M. (1995). Empirical generalizations about market evolution and stationarity. *Marketing Science, 14*(3), G109-G121.

Dekimpe, Marnik G., & Hanssens, Dominique M. (2011). The hidden powers of advertising investments. *Liber Amicorum of Peter Leeflang,* 209-223.

Deleersnyder, Barbara, Dekimpe, Marnik G., Sarvary, Miklos, & Parker, Philip M. (2004). Weathering tight economic times: the sales evolution of consumer durables over the business cycle. *Quantitative Marketing and Economics*, *2*(4), 347-383.

Deleersnyder, Barbara, Dekimpe, Marnik G., Steenkamp, Jan-Benedict E.M., & Leeflang, Peter S.H. (2009). The role of national culture in advertising's sensitivity to business cycles: an investigation across continents. *Journal of Marketing Research, 46*(5), 623-636.

Edell, Julie A., & Keller, Kevin L. (1989). The information processing of coordinated media campaigns. *Journal of Marketing Research*, *26*(2), 149-163.

Ewing, Michael T., du Plessis, Erik, & Foster, Charles (2001). Cinema advertising re-considered. *Journal of Advertising Research, 41*(1), 78-85.

Fagley, Nancy S. (1985). Applied statistical power analysis and the interpretation of nonsignificant results by research consumers. *Journal of Consulting Psychology*, *32*(2), 391-396.

Field, Chris A., & Welsh, Alan H. (2007). Bootstrapping clustered data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(3), 369-390.

Fisher, Marc, Albers, Sönke, Wagner, Niks, & Frie, Monica (2011). Dynamic marketing budget allocation across countries, products, and marketing activities. *Marketing Science*, *30*(4), 568-585.

Fok, Dennis, Horváth, Csilla, Paap, Richard, & Franses, Philip H. (2006). A hierarchical Bayes error-correction model to explain dynamic effects of price changes. *Journal of Marketing Research*, *43*(3), 443-462.

Fok, Dennis, Paap, Richard, & Franses, Philip H. (2003). Modeling dynamic effects of the marketing mix on market shares. *ERIM Report Series Research in Management*, ERS-2003-044-MKT.

Greene, William H. (2000). *Econometric analysis*, New Jersey: Pearson Education Inc.

Hair, Joseph F. Jr., Anderson, Rolph E., Tatham, Ronald L., & Black, William C. (1995). *Multivariate data analysis*, New York: Macmillan.

Hanssens, Dominique M., Parsons, Leonard J., & Schultz, Randall L. (2001). *Market response models : econometric and time series analysis*, Massachusetts: Springer.

Horváth, Csilla, & Fok, Dennis (2013). Moderating factors of immediate, gross, and net cross-brand effects of price promotions. *Marketing Science*, *32*(1), 127-152.

Hubbard, Raymond, & Armstrong, Scott (1992). Are null results becoming an endangered species in marketing? *Marketing Letters*, *3*(2), 127-136.

Joshi, Amit, & Hanssens, Dominique M. (2010). The direct and indirect effects of advertising spending on firm value. *Journal of Marketing*, *74*(1), 20-33.

Lamey, Lien, Deleersnyder Barbara, Dekimpe, Marnik G., & Steenkamp, Jan-Benedict E.M. (2007). How business cycles contribute to private-label success: evidence from the United States and Europe. *Journal of Marketing*, *71*(1), 1-15.

Lamey, Lien, Deleersnyder, Barbara, Steenkamp, Jan-Benedict E.M., & Dekimpe, Marnik G. (2012). The effect of business-cycle fluctuations on private-label share: What has marketing conduct got to do with it? *Journal of Marketing*, *76*(1), 1-19.

Leeflang, Peter S.H., & Wittink, Dick R. (1996). Competitive reaction versus consumer response: Do managers overreact? *International Journal of Research in Marketing, 13*(2), 103-119.

Leeflang, Peter S.H., Wittink, Dick R., Wedel, Michel, & Naert, Philippe A. (2000). *Building models for marketing decisions*, Dordrecht: Kluwer.

Levin, Andrew, Lin, Chien-Fu, & Chu, Chia-Shang J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, *108*(1), 1-24.

Lodish, Leonard M., Abraham, Magid, Kalmenson, Stuart, Livelsberger, Jeanne, Lubetkin, Beth, Richardson, Bruce, & Stevens, Mary Ellen (1995). How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *Journal of Marketing Research, 32*(2), 125–139.

Morrison, Donald G. (1969). On the interpretation of discriminant analysis. *Journal of Marketing Research*, *6*(2), 156-163.

Naert, Philippe A., & Leeflang, Peter S.H. (1978). *Building implementable marketing models*, New York City: Springer.

Naik, Prasad A. (2007). Integrated marketing communications: provenance, practice and principles. In Gerard J. Tellis, & Tim Ambler (Eds.), *The SAGE Handbook of Advertising* (pp. 35-53). London: Sage Publications.

Naik, Prasad A., & Peters, Kay (2009). A hierarchical marketing communications model of online and offline media synergies. *Journal of Interactive Marketing, 23*(4), 288-299.

Naik, Prasad A., Peters, Kay, & Raman, Kaylan (2008). Heterogeneous response functions in advertising. *Arbeitspapiere des Lehrstuhls für Innovation, Neue Medien und Marketing*.

Naik, Prasad A., & Raman, Kaylan (2003). Understanding the impact of synergy in multimedia communications. *Journal of Marketing Research*, *40*(4), 375-388.

Nijs, Vincent R., Srinivasan, Shuba, & Pauwels, Koen (2007). Retail-price drivers and retailer profits. *Marketing Science*, *26*(4), 473-487.

Pauwels, Koen, Erguncu, Selin, & Yildirim, Gokhan (2013). Winning hearts, minds and sales: how marketing communication enters the purchase process in emerging and mature markets. *International Journal of Research in Marketing*, *30*(1), 57-68.

Phillips, Peter C.B., & Perron, Pierre (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-46.

Rosenthal, Robert (1991). *Meta-analytic procedures for social research,* Newbury Park: Sage Publications.

Sawyer, Alan G., & Peter, J. Paul (1983). The significance of statistical significance tests in marketing research. *Journal of Marketing Research*, *20*(2), 122-133.

Saxonhouse, Gary R. (1976). Estimated parameters as dependent variables. *The American Economic Review*, *66*(1), 178-183.

Sethuraman, Raj, Tellis, Gerard J., & Briesch, Richard A. (2011). How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research*, *48*(3), 457-471.

Sheehan, Kim B., & Doherty, Caitlin (2001). Re-weaving the web: integrating print and online communications. *Journal of Interactive Marketing*, *15*(2), 47-59.

Srinivasan, Shuba, Vanhuele, Marc, & Pauwels, Koen (2010). Mind-set metrics in market response models: an integrative approach. *Journal of Marketing Research*, *47*(4), 672-684.

Steenkamp, Jan-Benedict E.M., & Geyskens, Inge (2014). Manufacturer and retailer strategies to impact store brand share: global integration, local adaptation, and worldwide learning. *Marketing Science*, *33*(1), 6-26.

Steenkamp, Jan-Benedict E.M., Nijs, Vincent R., Hanssens, Dominique M., & Dekimpe, Marnik G. (2005). Competitive reactions to advertising and promotion attacks. *Marketing Science*, *24*(1), 35-54.

Stephen, Andrew T., & Galak, Jeff (2012). The effects of traditional and social earned media on sales: a study of a micro lending marketplace. *Journal of Marketing Research*, *49*(5), 624-639.

Strube, Michael J. (1985). Combining and comparing significance levels from nonindependent hypothesis tests. *Psychological Bulletin*, *97*(2), 334-341.

Van Heerde, Harald J., Gijsenberg, Maarten J., Dekimpe, Marnik G., & Steenkamp, Jan-Benedict E.M. (2013). Price and advertising effectiveness over the business cycle. *Journal of Marketing Research*, *50*(2), 177-193.

Van Heerde, Harald J., Helsen, Kristiaan, & Dekimpe, Marnik G. (2007). The impact of a product-harm crisis on marketing effectiveness. *Marketing Science*, *26*(2), 230-245.

Yeung, Matthew C.H., Ramasamy, Bala, Chen, Junsong, & Paliwoda, Stan (2013). Customer satisfaction and consumer expenditure in selected European countries. *International Journal of Research in Marketing*, *30*(4), 406-416.

Table 1: Data coverage

Product class	*Number of categories / brands*	*Examples of included categories*	*Examples of included brands*
Food	55 / 132	Biscuits and cookies	Delacre
		Canned vegetables	Bonduelle
		Yoghurt	Danone
Beverages	16 / 50	Beers	Stella Artois
		Carbonated soft drinks	Fanta
		Mineral water	Evian
Household care	13 / 40	Air refreshing products	Febreze
		General cleaners	Mr. Propre
		Fabrics cleaning	Dash
Personal care	10 / 31	Deodorants	Axe
		Depilatory & shaving	Gillette
		Haircare	Head & Shoulders
Pet food	2 / 8	Dog & cat food	Sheba
		Other pet food	Vitakraft
Total	96 / 261		

Table 2: Media usage

Medium	*Adspend share sample-wide*	*Number of users*	*Adspend share among users*
TV	82.9%	231	83.6%
Radio	2.6%	81	5.7%
Newspaper	1.9%	123	3.1%
Magazine	3.6%	185	4.4%
Billboard	7.2%	117	11.9%
Cinema	1.8%	43	6.9%

Table 3: Number of media used

Number of media used	*Number of brands*
1	52
2	57
3	60
4	44
5	30
6	18
Total	261

Table 4: Brand-specific findings

Medium	*Number of users*	*Share of significant*[(a)] *positive short-run elasticities* (β_j^{ic})	*Share of significant*[(a)] *positive long-run elasticities* (γ_j^{ic})
TV	225	19.6%	28.0%
Radio	76	14.5%	18.4%
Newspaper	117	12.0%	12.8%
Magazine	181	9.9%	15.5%
Billboard	116	16.4%	12.9%
Cinema	40	7.5%	15.0%

(a) One-sided, 10% level.

Table 5: Meta-analytic findings

		Without self-selection correction				*With self-selection correction*			
		Short-run elasticity (β_k)		*Long-run elasticity (γ_k)*		*Short-run elasticity (β_k)*		*Long-run elasticity (γ_k)*	
	Number of contributing brands	*Estimate*[(a)]	*One-sided p-value*	*Estimate*[(a)]	*One-sided p-value*	*Estimate*[(a)]	*One-sided p-value*	*Estimate*[(a)]	*One-sided p-value*
TV	225	**0.0037**	0.000	**0.0070**	0.000	**0.0026**	0.002	**0.0043**	0.000
Radio	76	**0.0042**	0.002	**0.0045**	0.019	-0.0001	0.513	0.0024	0.249
Newspaper	117	**0.0028**	0.034	0.0009	0.356	0.0006	0.372	-0.0022	0.759
Magazine	181	**0.0033**	0.002	**0.0042**	0.000	**0.0024**	0.040	**0.0041**	0.002
Billboard	116	**0.0027**	0.020	**0.0037**	0.049	0.0017	0.152	0.0004	0.451
Cinema	40	0.0017	0.269	0.0017	0.245	0.0028	0.317	0.0059	0.148

(a) Effects significant at $p < 0.10$ (one-sided) are put in bold.

Table 6a: Robustness checks on the short-run meta-analytic results

	Focal model	*Alternative performance metric: market share*	*Alternative model specification: APAM(a)*	*Alternative DRS(b)*			*Stability*	
				η = 0.4	*η = 0.5*	*η = 0.6*	*First 50% of the data*	*Last 50% of the data*
TV	**0.0026**	**0.0009**	**0.0025**	**358.64**	**99.48**	**27.58**	**0.0023**	**0.0024**
Radio	-0.0001	0.0007	-0.0022	-468.19	-160.24	-52.63	-0.0032	0.0023
Newspaper	0.0006	-0.0006	-0.0013	-604.93	-165.06	-41.15	-0.0050	-0.0028
Magazine	**0.0024**	0.0003	**0.0019**	**581.47**	**197.10**	**64.14**	0.0023	**0.0058**
Billboard	0.0017	0.0003	0.0004	330.60	82.89	24.39	0.0077	0.0033
Cinema	0.0028	0.0017	0.0014	-786.19	-292.04	-89.69	-0.0004	-0.0024

Effects significant at $p < 0.10$ (one-sided) are put in bold.
(a) APAM = autoregressive partial-adjustment model.
(b) DRS = diminishing returns to scale, with η denoting the according rate, i.e. $\text{Advertising}^{\eta}$.

Table 6b: Robustness checks on the long-run meta-analytic results

	Focal model	*Alternative performance metric: market share*	*Alternative model specification: APAM*[(a)]	*Alternative DRS*[(b)]			*Stability*	
				$\eta = 0.4$	$\eta = 0.5$	$\eta = 0.6$	*First 50% of the data*	*Last 50% of the data*
TV	**0.0043**	**0.0017**	**0.0191**	**554.15**	**155.25**	**42.68**	**0.0036**	**0.0055**
Radio	0.0024	0.0004	-0.0282	-605.71	-223.02	-74.15	-0.0011	0.0047
Newspaper	-0.0022	-0.0021	0.0117	-846.13	-220.28	-53.32	-0.0081	-0.0039
Magazine	**0.0041**	0.0009	0.0059	**992.53**	**329.15**	**106.32**	**0.0054**	**0.0058**
Billboard	0.0004	-0.0014	-0.0411	540.85	147.45	42.79	0.0124	0.0069
Cinema	0.0059	0.0023	-0.0092	526.73	120.01	34.96	0.0011	0.0031

Effects significant at $p < 0.10$ (one-sided) are put in bold.
(a) APAM = autoregressive partial-adjustment model.
(b) DRS = diminishing returns to scale, with η denoting the according rate, i.e. $\text{Advertising}^{\eta}$.

Table 7: Tests for synergy effects with the small media

Synergy between	*Number of common users*[(a)]	*Share (%) of significant*[(b)] *positive estimates (short | long run)*	*Meta-analytic estimate*[(c)] *(short | long run)*	*Meta-analytic one-sided p-value (short | long run)*
Billboard & TV	41	7 | 15	-0.0005 | 0.0005	0.998 | 0.232
Billboard & Radio	24	8 | 17	-0.0010 | **0.0045**	0.919 | 0.011
Billboard & Newspaper	31	6 | 10	-0.0005 | 0.0011	0.953 | 0.161
Billboard & Magazine	42	5 | 12	-0.0001 | 0.0006	0.579 | 0.304
Billboard & Cinema	13	8 | 8	0.0009 | 0.0024	0.231 | 0.369
Cinema & TV	16	0 | 13	0.0008 | -0.0019	0.191 | 0.776
Cinema & Radio	8	25 | 13	0.0018 | -0.0067	0.198 | 0.982
Cinema & Newspaper	9	22 | 22	-0.0010 | 0.0026	0.666 | 0.254
Cinema & Magazine	12	0 | 0	0.0007 | 0.0025	0.347 | 0.236

(a) This number only includes those brands with VIF values smaller than 10.
(b) One-sided, 10% level.
(c) Effects significant at $p < 0.10$ (one-sided) are put in bold.